Media3.332

N° d'ordre :
N° de série :

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
Ministry of Higher Education and Scientific Research

ECHAHID HAMMA LAKHDAR UNIVERSITY - EL OUED
FACULTY OF EXACT SCIENCES
Computer Science department

End of Study Memory

Presented for the Diploma of

ACADEMIC MASTER

Domain : **Mathematics and Computer Science**
spinneret: **Computer Science**
Speciality : **Artificial Intelligence and Distributed Systems**

Presented by :

- **Nacira Berrehouma**
- **Ridha Berrehouma**

Theme

Deep Learning and Big Data Architecture for Seismic Data Analyses

Supported in: - - 2019 In front of jury:

M.	Kholladi Najoua	President
M.	Abdelkader Laouid	Reporter
Dr.	Mohammed Anouar Naoui	Supervisor

Année Universitaire: 2018/2019

ABSTRACT

Deep learning algorithms have gained a lot of interest in recent years by obtaining state-of-the art results in various problems arising in the fields of computer vision, automatic speech recognition and natural language processing, this project aims to build a deep convolution neural network model for classify a 2D slices of seismic volumetric reflection data as either "salt" or "not salt", applying big data to train the classifier obtained , The result indicate that the accuracy of the prediction of image's class is 68% ,

key words:deep learning, big data, convolution neural network, salt interpretation

DEDICATION

We dedicate this modest work to whom brought us love, support and happiness , to the dearest people to our hearts, To our dear parents, for their; love, sacrifice, patience, moral and material support from our childhood until this day, To all our brothers and sisters , who supported us, encouraged us and who took the way with us along all the courses of our studies, To all our dear nephews , to all our uncles and cousins and everyone in our big family. To our dear friends

Nacira & Ridha

KNOWLEDGMENT

“ Thanks to the one God, Light of the heavens and the earth, who help and guide ”

we must first of all thank “ ALLAH ” the almighty, who gave us the power, the will and the patience to develop this work .
Our sincerest thanks to our mentor Mr. Naoui Mohammed Anouar for his availability, his contributions, his valuable guidance and his understanding throughout the development of this memoir.
We would also like to thank the members of the jury very much for accepting, examine and evaluate this modest work .
We also thank all our teachers of Licence and Master in the department of Computing Science.

Thank you.

Table of Contents

List of Figures

INTRODUCTION

Object presentation

Drilling for oil and gas is an expensive process and takes considerable time, companies in the oil and gas industry invest millions of dollars in an effort to improve their understanding of components of the sub-surface, an integral part of this effort consists of using traditional workflows for interpreting large volumes of seismic data. Geo-scientists are required to manually define relationships between geological features and seismic patterns. Therefore, it is necessary to find an alternative way, so geologists and oil and gas industry companies resorted to a seismic survey, where seismic waves give a lot of information about what is inside the earth without need to dig.

In this erea, deep learning become foundational to the digital transformation of the oil and gas industry and they have a dramatic impact on the exploration and production of hydrocarbons, where it can resolves two significant issues in seismic interpretation:

1. The 'Big Data' problem of trying to interpret dozens, if not hundreds, of volumes of data.

2. The fact that humans cannot understand the relationship of more than three types of data all at once.

The problematic

The main problematic of this thesis concerns the identification of salt layers of a seismic image by a computer which often coexist with gas and oil under the ground by proposing a Deep Learning and Big Data architecture for seismic analysis. To do so, we will follow the work plan described in the next section.

Work plan

This thesis contains four chapters as following :

- **Chapter one**

  in this chapter we will identify some of the geological components contained in the layers of the earth and we will look at the way to obtain a seismic image that can be used by geologists in the latter of this chapter we will take a view on the last work done by the researchers in the field of seismic analyses and computer vision.

- **Chapter tow**

  this chapter talk about the theoretical background of different algorithms of machine learning, in the last we show clearly how machine learning and deep learning relate to each other

- **Chapter three**

  this chapter focus on the use of Deep Learning in Big Data Analytic , the advantage of using the big data with deep learning and the challenges facing that tow techniques.

- **Chapter four**

  in the last chapter we propose our convolution neural network model and we will give a comparative study between another model used in seismic analysis.

Finally, we achieve this thesis by a general conclusion where we resume our main contribution and give some orientation for further work.

CHAPTER 1

SEISMIC SURVEY AND SALT STRUCTURES

1.1 Introduction

1.1.1 Seismic Survey

The *Seismic Survey* is an essential part of the whole cycle of petroleum exploration and production, it can be applied either on offshore (marine) or onshore (land) by creating an artificial waves source using explosive charge such as air-guns for offshore exploration and dynamite or specialized trucks for onshore exploration, which contains a heavy plate placed on the surface of the ground that is vibrated to generate waves which are bounced off underground rock formations and the waves that reflect back to the surface are captured by seismic receivers (hydrophones and geophones).

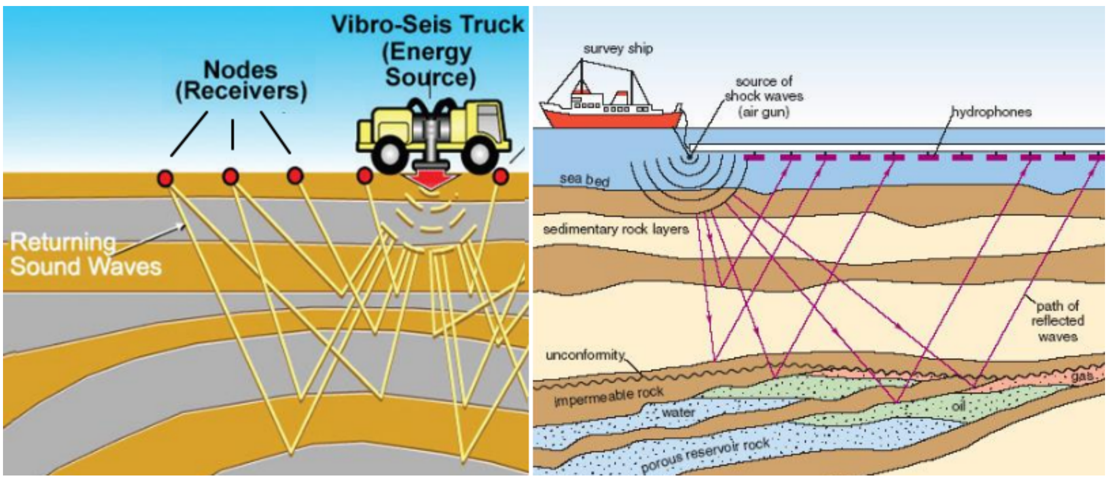


Figure 1.1: seismic survey

The hydrophone is a device designed for use in detecting seismic energy in the form of pressure changes in water during marine seismic acquisition, The geophone is a device used in surface seismic acquisition

The geophone is a device used in surface seis- mic acquisition, both onshore and on the seabed offshore, that detects ground velocity produced by seismic waves and transforms the motion into electrical impulses.

The time taken to travel from the source to the receivers can indicate features such as rock density and the likely presence of fluids or gases. This can help build an image of the subsurface [1]

Figure 1.2: Geophone

Figure 1.3: Hydrophone

1.1.2 Salt Structures

The *Salt Structures* present in the subsurface are a widespread geological phenomenon, it is formed as a result of the difference in density of salt and the layer above it, the salt is less dense then it pushes toward the top forming dome in sedimentary layers above it, if it contains oil, it moves towards the outer sides of the salt dome and is confined between the sedimentary layers, and thus constituted exploitable oil reservoirs.

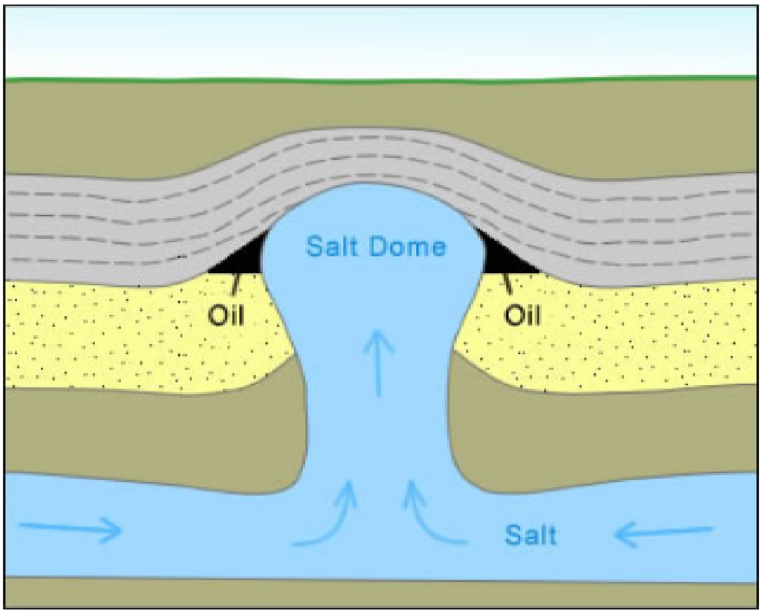


Figure 1.4: Salt dome

Due to the physical properties of salt, identifing the salt structure is consider as one of challenges of seismic imaging, where the salt density is usually 2.14 g/cc which is lower than most surrounding rocks. The seismic velocity of salt is 4.5 km/sec[2], which is usually faster than its surrounding rocks. This difference creates a sharp reflection at the salt-sediment interface. Usually salt is an amorphous rock without much internal structure. This means that there is typically not much reflectivity inside the salt, unless there are sediments trapped inside it. The unusually high seismic velocity of salt can create problems with seismic imaging.

1.1.3 Why Interpret Salt ?

There are at least three reasons for salt interpretation:

- First, salt is an effective seal, so it forms the boundaries of many hydrocarbon traps. These traps cannot be mapped properly without interpreting the salt contact.
- Second, salt structures can profoundly influence sediment transport , and are consequently major controls on reservoir distribution, understanding salt geometry and evolution is therefore essential in predicting reservoir location in a salt basin.
- Finally, salt's unique physical properties create challenges for subsalt and salt-flank imaging in many basins, a principal technique for solving these challenges is prestack depth migration , which requires an accurate salt model. A large part of modern salt interpretation is thus directed at velocity-model building for prestack depth migration.

1.2 Seismic Data Analysis

Seismic Data Analysis is a main process to extract valuable information. this process flow the next steps:

1.2.1 Seismic Data Acquisition

The first step in Seismic data analysis is to acquire data, which in most cases is carried out from the surface, as explained previously, this step consists of using adequate devices (wave generator, hydrophone, geophone, etc.) to construct the base of data to be treated. At the end of this stage we get set of a noisy seismic images

1.2.2 Processing of Seismic Data

Seismic technology has achieved amazing feats in exploration and production activities in the past few decades. What we record in the acquisition stage is called raw seismic data which contains real signals together with noise and multiples.

Next this raw data must be processed, where the raw data is put through many complex procedures using powerful computers to get better images of the subsurface. The prime objective in the processing stage is to enhance the signal and suppress the coherent and non coherent noises and multiples [1]

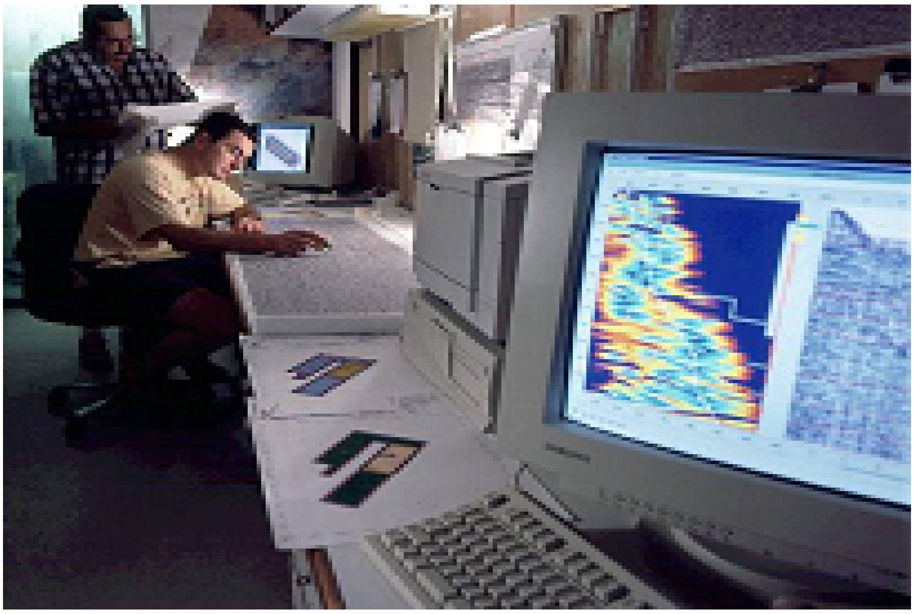


Figure 1.5: Seismic processors and powerful computer workstations

1.2.3 Inversion of Seismic Data

In a Seismic Inversion, the original reflective data, as typically, recorded routinely, then is converted from a reflection signal to a rock property known as impedance, which itself is the multiplication of sonic velocity and bulk density .

In a conventional seismic reflectivity section the strong amplitudes are associated with the boundaries between geological formations, such as the top reservoir. This type of data is most suited to structural interpretation. In an inverted dataset, the amplitudes are now describing the internal rock properties, such as lithology type, porosity or the fluid type in the rocks (brine or hydrocarbons). Inverted data is ideal for stratigraphic interpretation and reservoir characterization. [3]

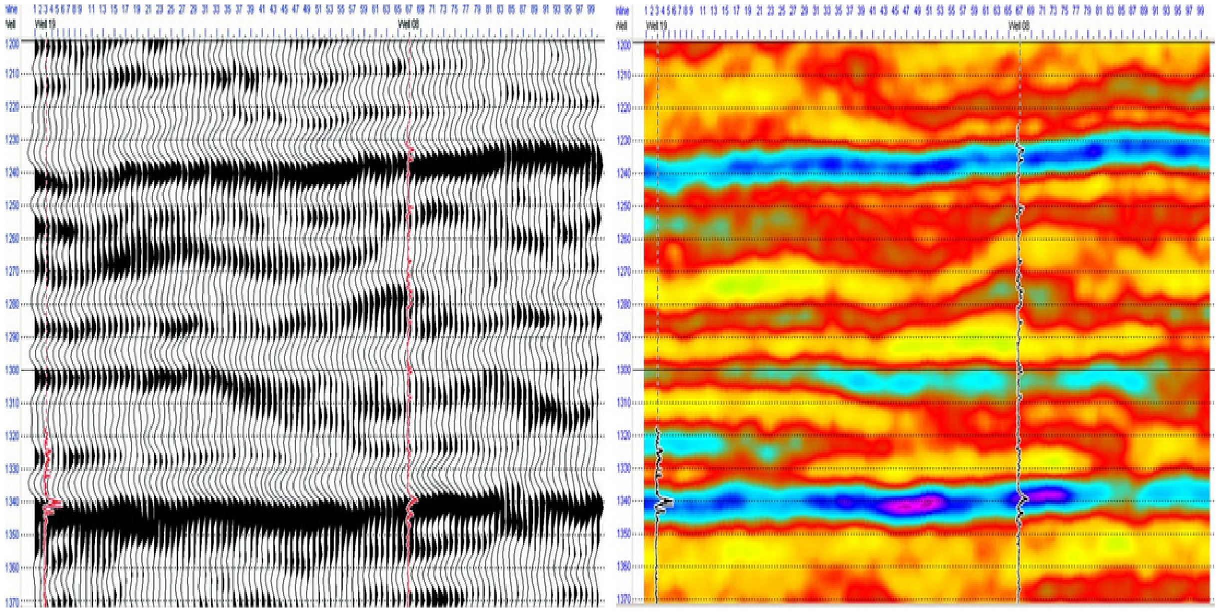


Figure 1.6: seismic reflection signal inverted to seismic rock property

The left of the Figure 1.6 shows seismic reflection signal and on the right the seismic rock properties, we note that the strong amplitudes are shown as the blue/purple colors in the right side.

1.2.4 Interpretation of Seismic Data

The goal of seismic interpretation is to make sense of a seismic image by obtain a coherent geological story from the map of processed seismic reflections[4], in another meaning it includes the analysis of the 2D or 3D data set obtained via the previously steps in order to produce structural maps that reflect the spatial variation in the depth of certain geological layers. Using these maps, hydrocarbon traps can be identified and models can be created from the bottom surface. However, the seismic dataset rarely gives a clear enough picture to do so. This is mainly due to seismic accuracy but often noise and handling difficulties also lead to lower image quality. Because of this, there is always a degree of uncertainty in seismic interpretation and a particular data set can contain more than one solution that fits the data. In such a case, more data would be needed to restrict the solution.

In hydrocarbon exploration, the features that the interpreter is particularly trying to delineate are the parts that make up a petroleum reservoir – the source rock, the reservoir rock, the seal and trap.

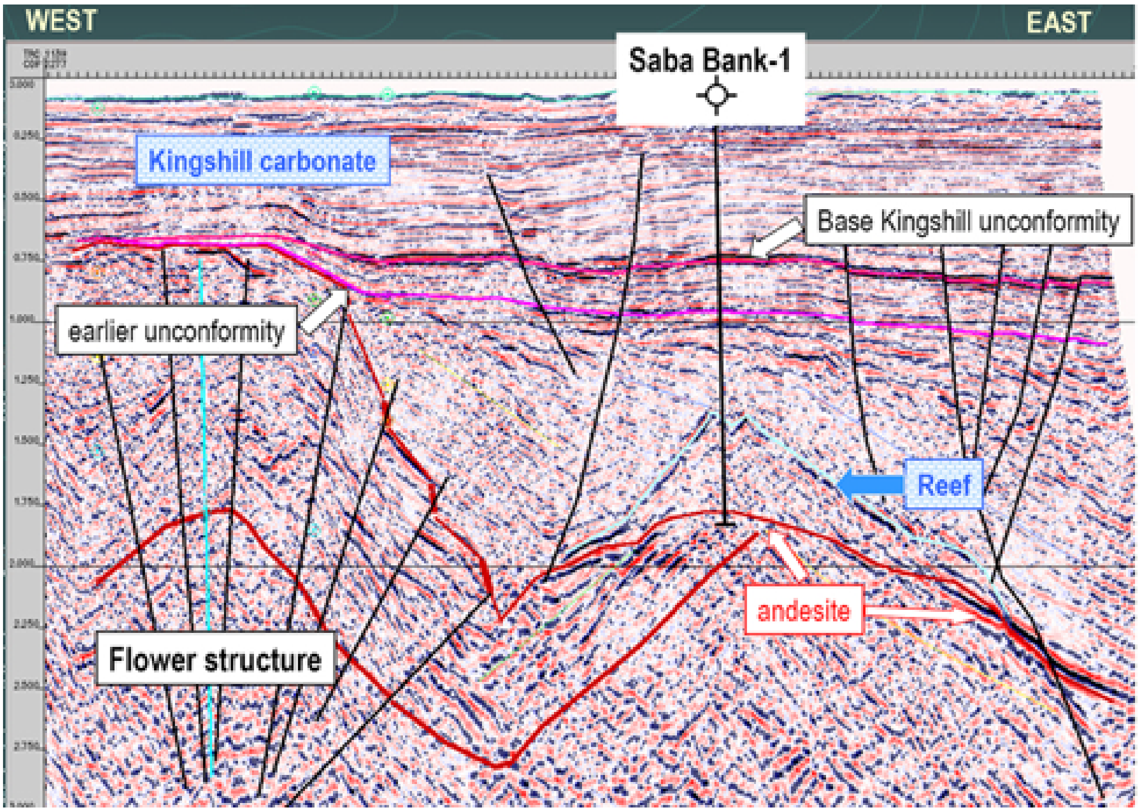


Figure 1.7: Seismic profile with interpretation of horizons and structures

1.3 State of the Art

In this section, we try to give some important results related to Seismic Data processing.

1.3.1 Seismic and Computer Vision

Specific Big Data domains including computer vision [5] and speech recognition [6], have seen the advantages of using Deep Learning to improve classification modeling results but, there are a few works on Deep Learning architecture for seismic analysis. In 2007 OlivierLavialle et al [7] proposed a pre-processing step based on a non-linear diffusion filtering leading to a better detection of seismic faults, the non-linear diffusion approaches are based on the definition of a partial differential equation that allows to simplify the images without blurring relevant details or discontinuities. In 2002, I. Pitas and C. Kotropoulos [8]. proposed tow methodes for the texture analysis and segmentation of geophysical images based on the detection of the seismic horizons and on the calculation of their features where the horizons are clustered into classes according to one or several of their features. The last year Mikhail Karchevskiy, Insaf Ashrapov and Leonid Kozinkin [9] present A Deep Learning approach for automatic salt deposits segmentation, demonstrating the great performance of several novel deep learning techniques merged into a single neural network which achieved the 27th place Kaggle competition for salt deposits segmentation problem in seismic image data. Using a U-Net with ResNeXt-50 encoder pre-trained on ImageNet as their base architecture

1.3.2 Software for Seismic Data Analysis

There are many software for geophysical data processing and interpretation, in following we list some software tools with a pref explanation

SeismicLab

SeismicLab is a MATLAB seismic data processing package. The package can be used to process small seismic data sets, and it is mainly intended for research and teaching purposes. it provide a scripts to read and write and Seismic Unix (SU) data (open source seismic utilities package which was supported by the Center for Wave Phenomena) . A particular feature of this package is that the SEGY headers are loaded into a structure that can be easily assessed by MATLAB.

ANSeismic

ANSeismic is a simple assistant program developed for implementing earthquake analyses of structures, with ANSYS and SAP2000. The seismic records are loaded from PEER (Pacific Earthquake Engineering Research Center) and earthquake analysis files are produced in ANSYS Parametric Design Language (APDL). Anyone who modeled a structure in ANSYS can use the analysis files produced with ANSeismic by just calling them. [10]

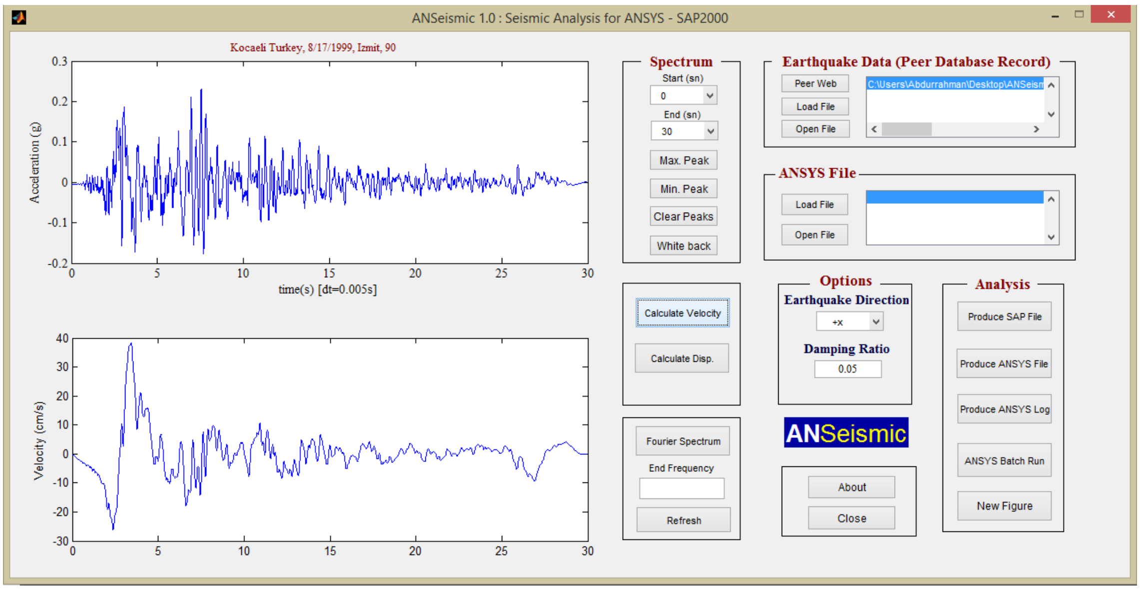


Figure 1.8: Main window of ANSeismic program

1.4 Conclusion

In this chapter. We have tried to feed the reader with the principle concepts necessary to deal with propose of this thesis. After giving general definitions, we walked out the steps of seismic date analysis. In the end of this chapter, we have highlighted the main contributions in this field. Next chapters focus on the using of big data and deep learning techniques.

CHAPTER 2

MACHINE LEARNING AND DEEP LEARNING

2.1 Introduction

Nowadays, both of *Machine Learning* (ML) and *Deep Learning* (DL) are getting very hot importance in the AI field, because these learning algorithms can be used in several fields such as software engineering, image processing, financial investment, etc. In this chapter we will talk about the theoretical background for different algorithms of machine learning, and in the last section, we will show clearly how machine learning and deep learning relate to each other.

2.2 Machine learning

Many definitions were given in the literature, we chose the definition of *M. D. Youcef*: "The ML is an application of artificial intelligence that allow computers to automatically learn -without any human intervention or assistance- from a set of observations or data that we call *learning set*".[11]

2.2.1 Types of Machine Learning

Machine learning algorithms are often categorized as following [12] :

- **Supervised machine learning algorithms :**

  Can apply what has been learned in the past to new data using labeled examples to predict future events. Starting from the analysis of a known training dataset, the learning algorithm produces an inferred function to make predictions about the output values. The system is able to provide targets for any new input after sufficient training. The learning algorithm can also compare its output with the correct, intended output and find errors in order to modify the model accordingly.

- **Unsupervised machine learning algorithms :**

  Are used when the information used to train is neither classified nor labeled. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data. The system doesn't figure out the right output, but it explores the data and can draw inferences from datasets to describe hidden structures from unlabeled data.

- **Semi-supervised machine learning algorithms :**

  Fall somewhere in between supervised and unsupervised learning, since they use both labeled and unlabeled data for training, typically a small amount of labeled data and a large amount of unlabeled data. The systems that use this method are able to considerably improve learning accuracy. Usually, semi-supervised learning is chosen when the acquired labeled data requires skilled and relevant resources in order to train it / learn from it. Otherwise, acquiring unlabeled data generally doesn't require additional resources.

- **Reinforcement machine learning algorithms :**

  Is a learning method that interacts with its environment by producing actions and discovers errors or rewards. Trial and error search and delayed reward are the most relevant characteristics of reinforcement learning. This method allows machines and software agents to automatically determine the ideal behavior within a specific context in order to maximize its performance. Simple reward feedback is required for the agent to learn which action is best; this is known as the reinforcement signal.

2.2.2 Algorithms of Machine Learning

Decision Trees

The decision tree is an algorithm based on a graph model (trees) to define the final decision. Each node has a condition, and the branches are based on this condition (True or False). The more we descend into the tree, the more we combine the conditions. The picture above illustrates this operation.[13]

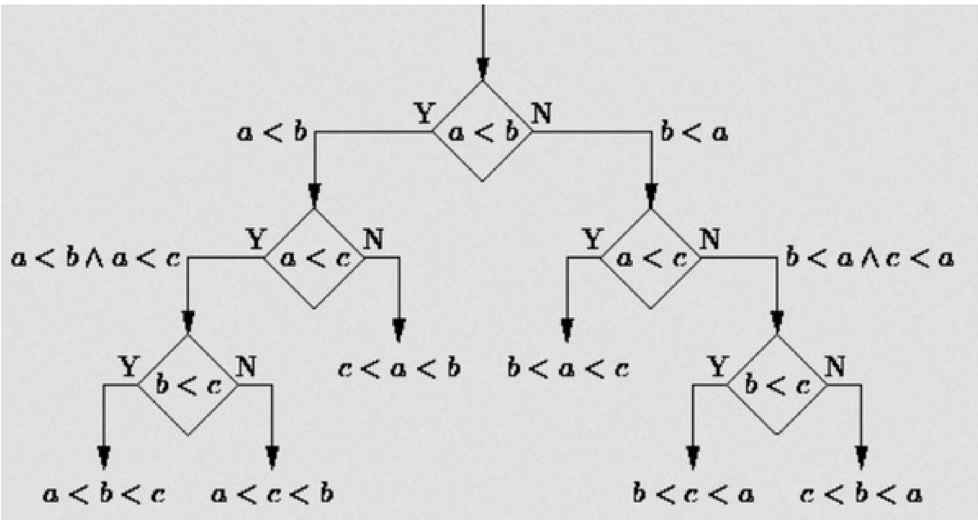


Figure 2.1: Decision Trees

Linear Regression

Linear regression algorithms model the relationship between predictor variables and a target variable. The relation is modeled by a mathematical function of prediction. The simplest case is univariate linear regression. She will find a function in the form of a right to estimate the relation. Multivariate linear regression occurs when several explanatory variables intervene in the prediction function. And finally, polynomial regression can model complex relationships that are not necessarily linear.[13]

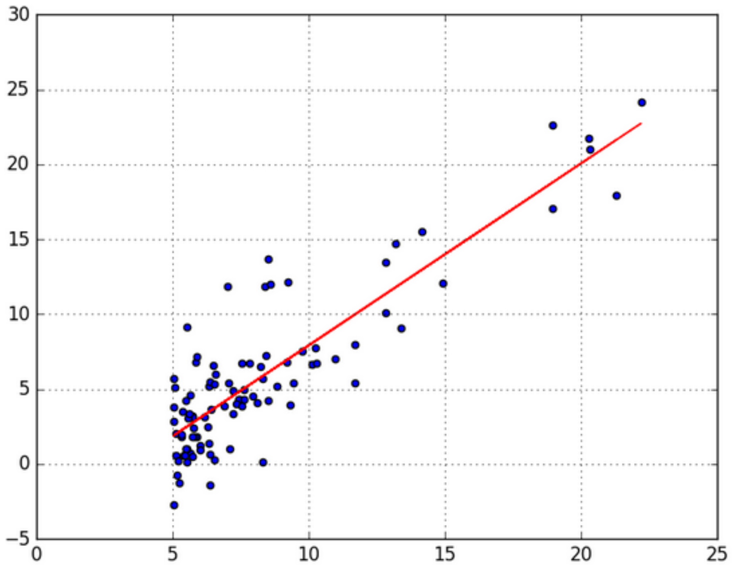


Figure 2.2: Linear Regression

K-Nearest Neighbors

K-Nearest Neighbors is a supervised classification algorithm. Each observation of the learning set is represented by a point in a space with n dimensions where n is the number of predictor variables, to predict the class of an observation, we look for the k nearest points of this example, The class of the target variable, is the one that is most represented among the k nearest neighbors. There are variants of the algorithm or we weight the k observations according to their distance to the example we want to classify [14].

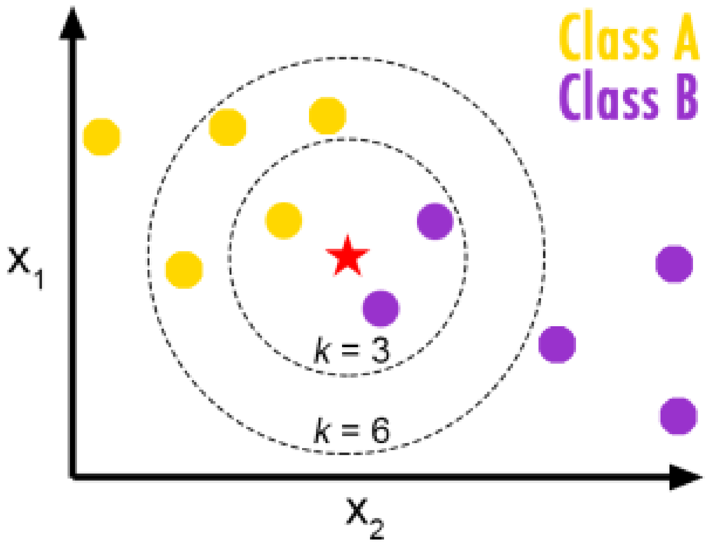


Figure 2.3: K-Nearest Neighbors

For $k = 3$ the majority class of the central point is class B, but if we change the value from neighborhood $k = 6$ the majority class becomes class A.

Naïve Bayes

Naïve Bayes is a fairly intuitive classifier to understand. It is based on Bayes' theorem of conditional probabilities. The picture above is the formula of Bayes' theorem. Naive Bayes assumes a strong (naive) assumption. Indeed, it assumes that the variables are independent of each other. This simplifies the calculation of probabilities. Generally, Naïve Bayes is used for text classifications (based on the number of word occurrences).[15]

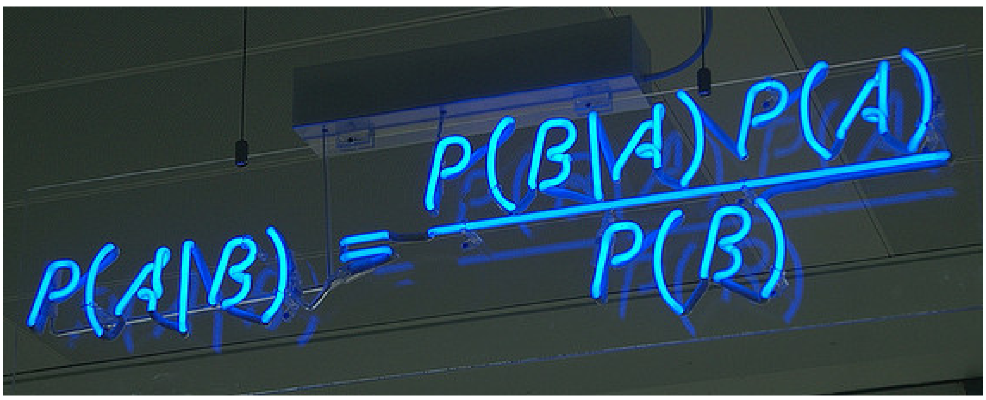


Figure 2.4: Bayes theorem formula

Support Vector Machine (SVM)

Support Vector Machine (SVM) is a binary classification algorithm. If we take the image below, we have two classes (Imagine that it's e-mails, and that Spam e-mails are red and non-spam are blue). the SVM will opt to separate the two classes by a line.

the SVM will choose the clearest separation possible between the two classes (like the green line). This is why it is also called Large Margins classifier (classifier at wide margins).[16]

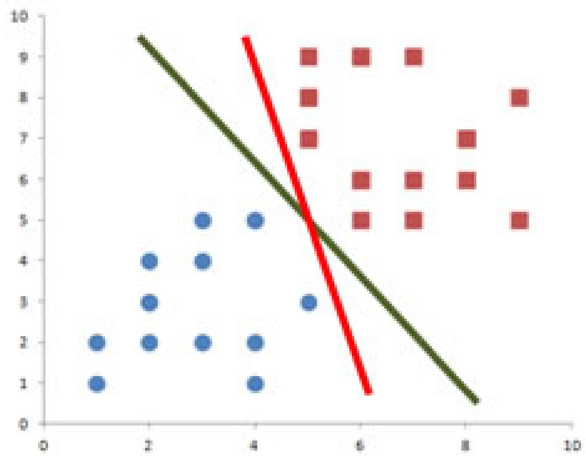


Figure 2.5: Support Vector Machine

Anomaly Detection

Anomaly Detection is a Machine Learning algorithm for detecting abnormal patterns. Imagine for example that you receive in your bank account 2000 monthly and that one day you deposit 10 000 at once. The algorithm will detect this as an anomaly. This algorithm is very useful for detecting fraud in banking transactions and intrusion detections.[17]

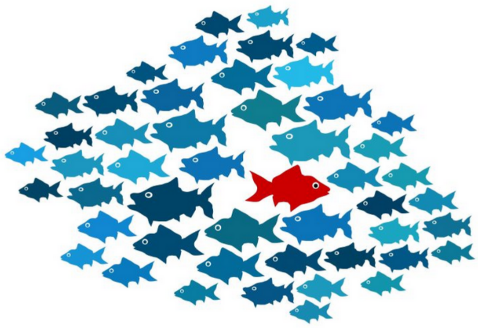


Figure 2.6: Anomaly Detection

K-Means

K-Means is a clustering algorithm in Unsupervised Learning. It is given a set of elements (data), and a number of groups K. K-means will segment into K groups the elements. Grouping is done by minimizing the Euclidean distance between the center of the cluster and a given element.[18]

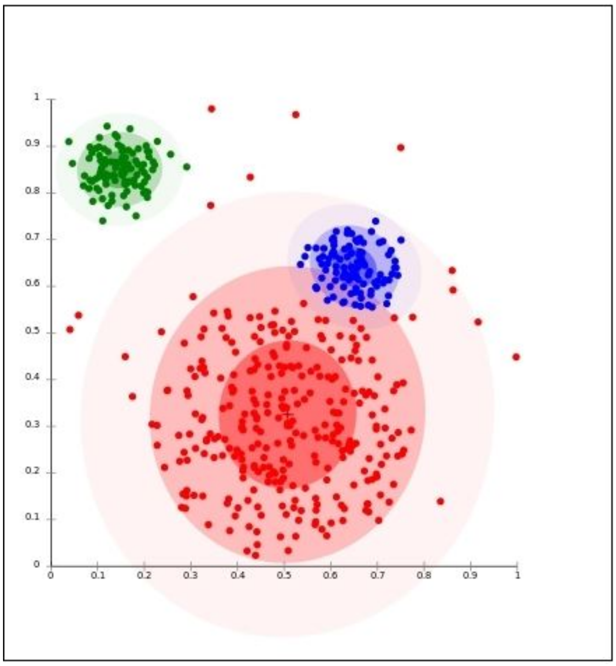


Figure 2.7: k means

Neural Network

Neural networks are inspired by the neurons of the human nervous system. They allow to find complex patterns in the data. These neural networks learn a specific task based on the training data.

these networks consist tree types of layer, Input Layer that receive input data, Hidden Layers receive the data from the precedent layer, Finally the Output Layer witch produces the classification result [19].

circles in the figure 2.8 represent one neuron witch is more detailed in the right side of figure 2.9.

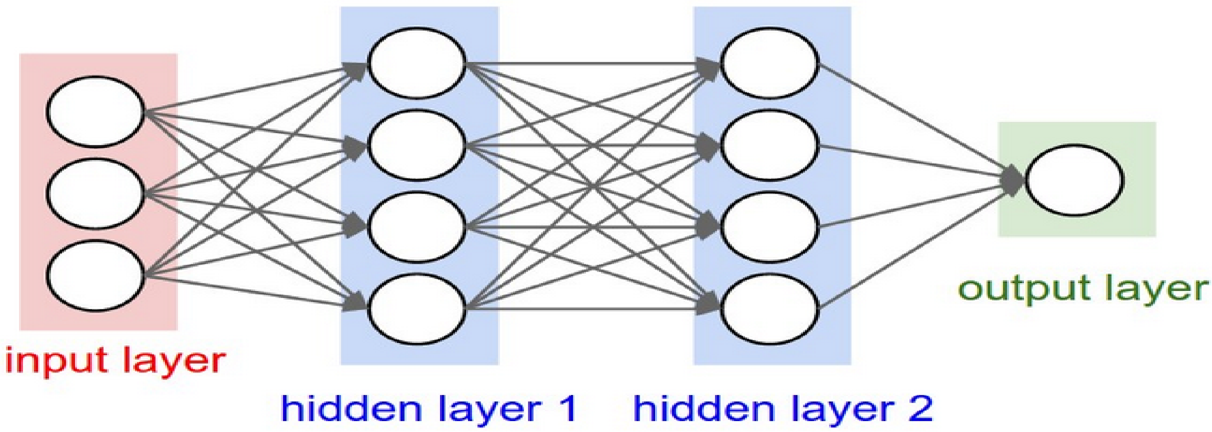


Figure 2.8: neuron network layers

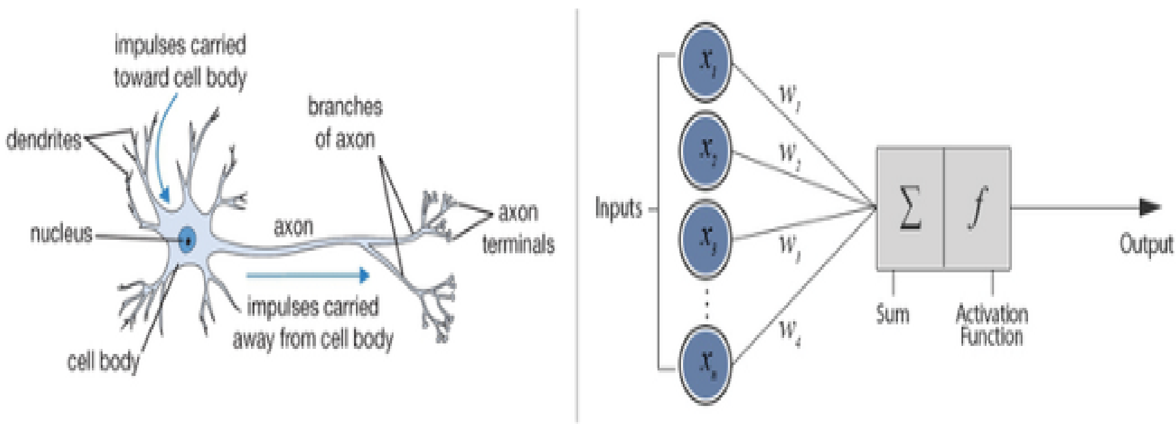


Figure 2.9: biological neuron VS artificial neuron

2.3 Deep Learning

The term "Deep Learning" was introduced for the first time at Machine Learning by Dechter (1986) and artificial neural networks by Aizenberg et al (2000) [20]

Deep Learning is a specific subset of machine learning, Deep learning algorithms are motivated by the field of artificial intelligence, which has the general goal of emulating the human brain's ability to observe, analyze, learn, and make decisions, especially for extremely complex problems. This algorithms was inspired by the structure and functioning of the brain.

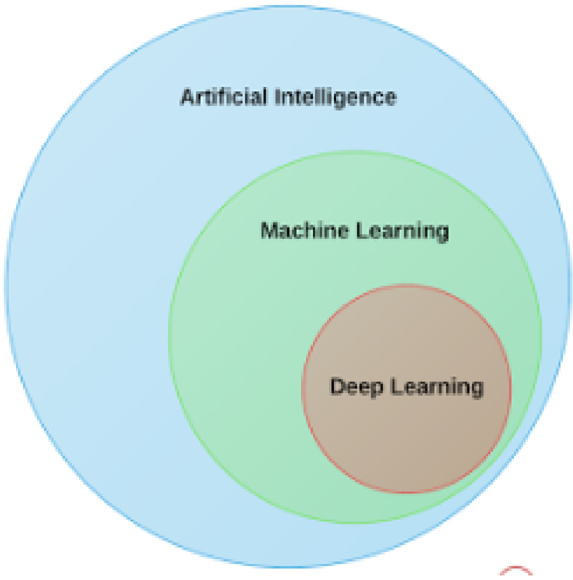


Figure 2.10: The relationship between artificial intelligence, Machine learning and deep learning

Deep Learning is based on the idea of artificial neural networks (ANN) which are layered structures inspired by the human brain. and is designed to handle large amounts of data by adding layers to the network. The term "deep" usually refers to the number of hidden layers in the neural network. Traditional neural networks only contain 2-3 hidden layers, while deep networks can have as many as 150.

To understand how Deep Learning works, we take a simple example, suppose we want to make a system that can recognize faces of different people in an image . If we solve this as a typical machine learning problem, we will define facial features such as eyes, nose, ears etc. and then, the system will identify which features are more important for which person on its own in opposite Deep Learning automatically finds out the features which are important for classification because of deep neural networks. The figure 2.11 explains this example .

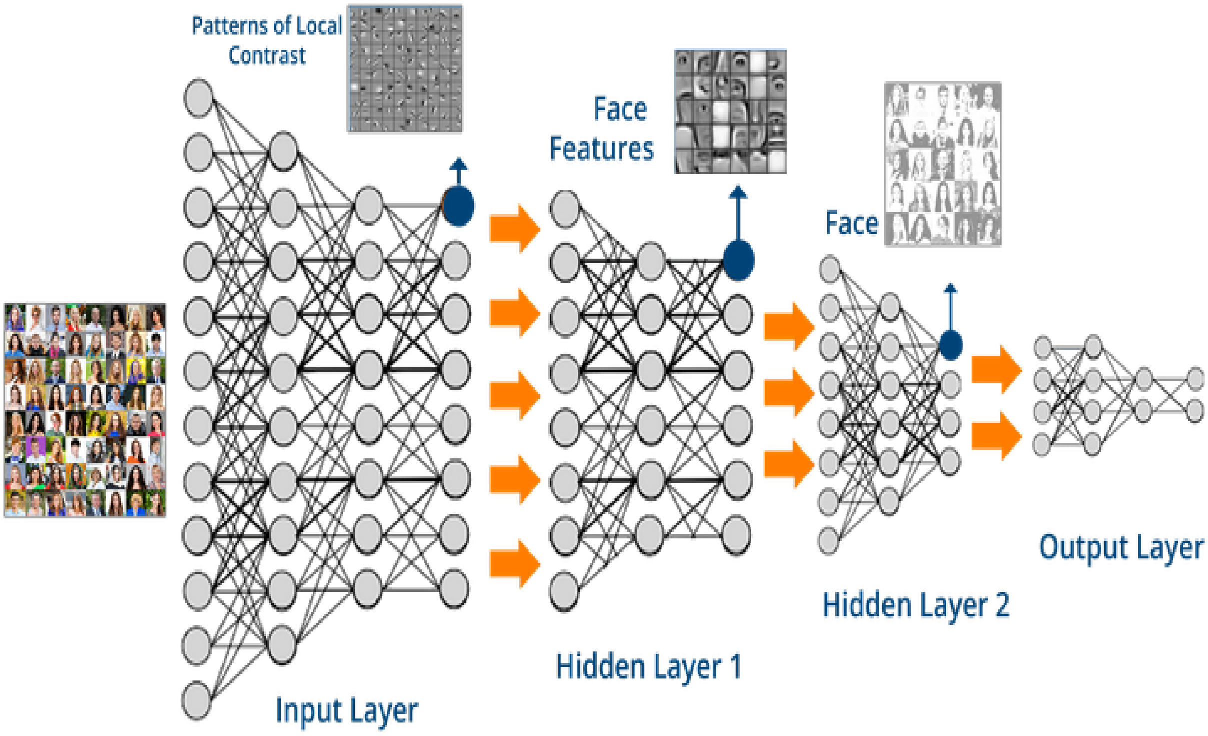


Figure 2.11: Face Recognition using Deep Networks

Convolutional Neural Networks (CNN) is one of the most widely used algorithms in deep learning. more details in next section

2.3.1 Convolution Neural Network

Convolution neural networks are considered as a good solution in classification tasks, in which a model learns to perform classification tasks directly from input data with minimal human intervention.
A convolutional neural network can have tens or hundreds of layers that each learn to detect different features of an image. Filters are applied to each training image at different resolutions, and the output of each convolved image is used as the input to the next layer. The filters can start as very simple features, such as brightness and edges, and increase in complexity to features that uniquely define the object.
The process of building a Convolution Neural Network always involves three major steps :

Convolutional Layers

A convolution is the simple application of a filter to an input that results in an activation. Repeated application of the same filter to an input results in a map of activations called a feature map, indicating the locations and strength of a detected feature in an input, such as an image. A convolutional layer can be created by specifying both the number of filters to learn and the fixed size of each filter, often called the kernel shape.

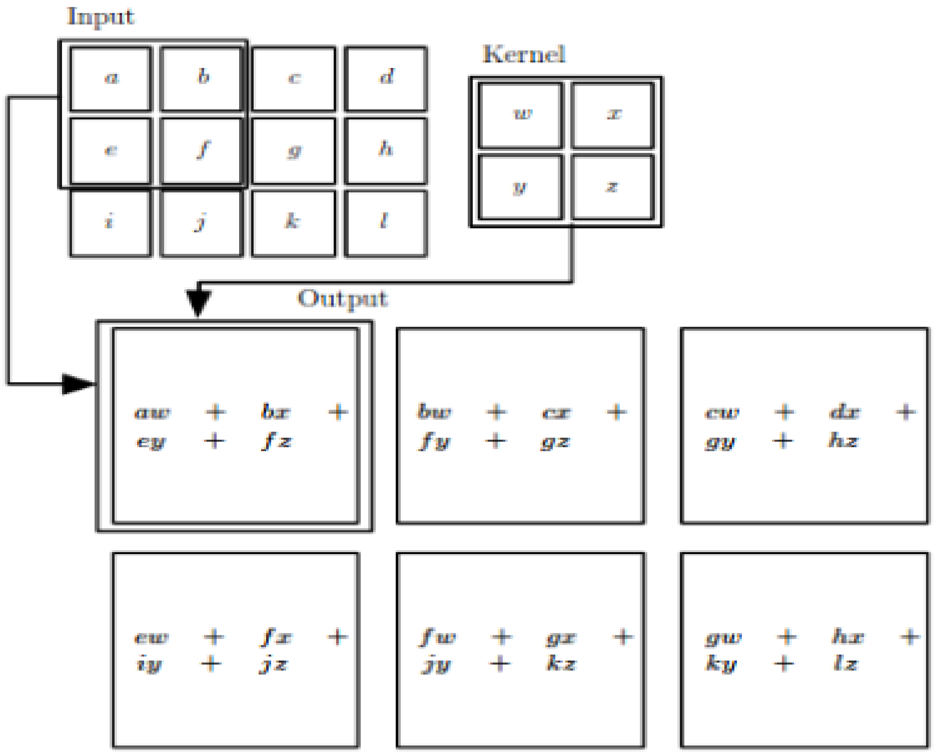
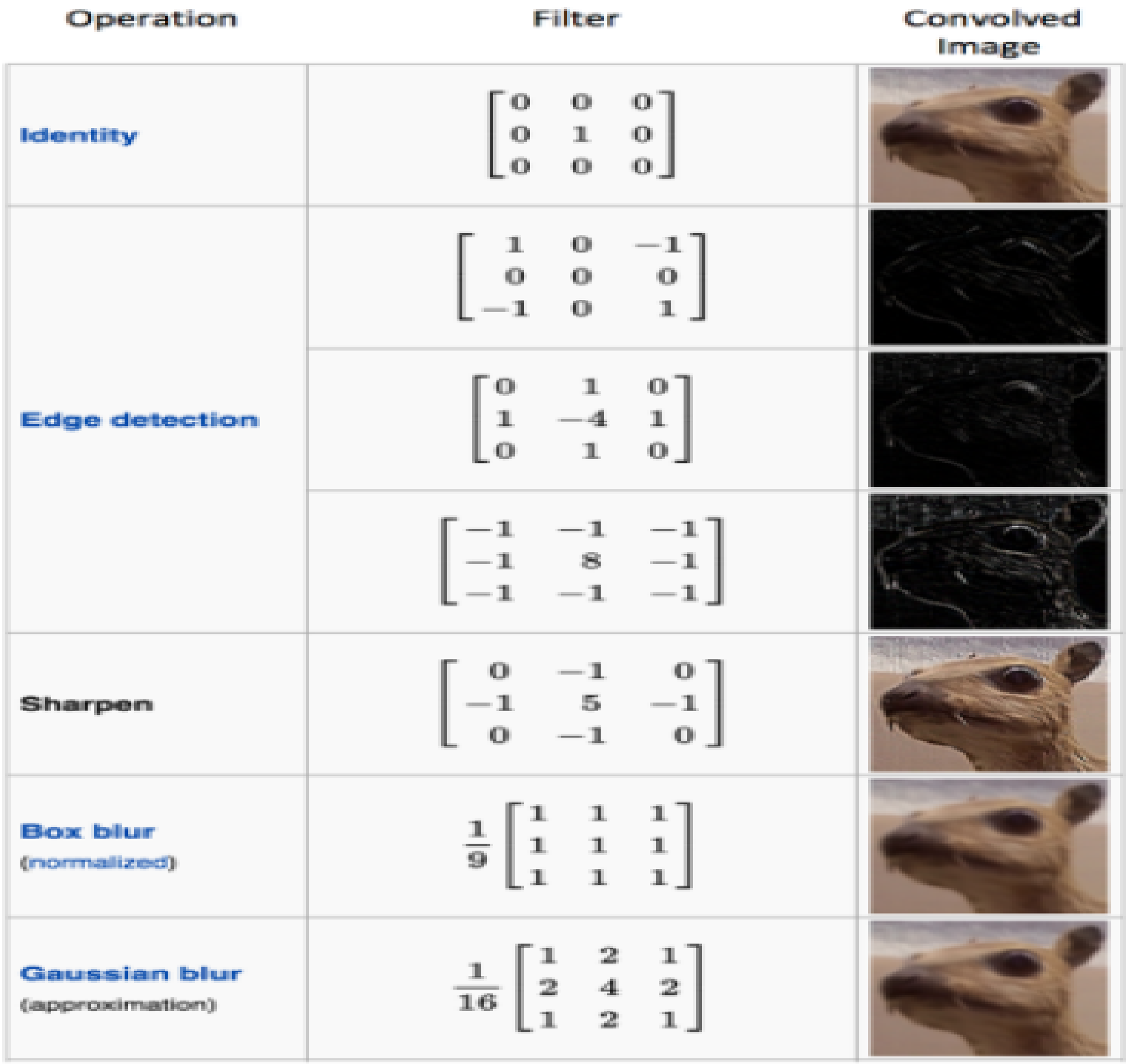


Figure 2.12: 2D convolution example

The figure 2.13 illustrate a different types of filters can be applied into an input image



Operation	Filter	Convolved Image
Identity	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	
Edge detection	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$	
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	
Sharpen	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
Box blur (normalized)	$\frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	
Gaussian blur (approximation)	$\frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$	

Figure 2.13: common filters

Pooling Layers

Pooling layers provide an approach to downsampling feature maps by summarizing the presence of features in patches of the feature map. Maximum pooling, or max pooling, is a pooling operation that calculates the maximum, or largest, value in each patch of each feature map.

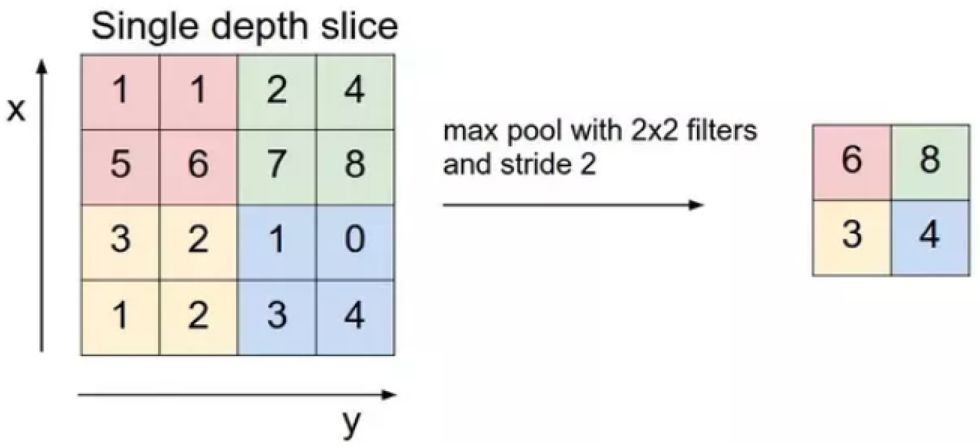


Figure 2.14: Max Pooling example

Classifier Layer

Once the features have been extracted, they can be interpreted and used to make a prediction, such as classifying the type of object in a photograph. This can be achieved by first flattening the two-dimensional feature maps, and then adding a fully connected output layer. For a binary classification problem, the output layer would have one node that would predict a value between 0 and 1 for the two classes.

famous convolution networks

LeNet :

LeNet-5, a pioneering 7-level convolutional network by LeCun et al in 1998, that classifies digits, was applied by several banks to recognise hand-written numbers on checks (cheques) digitized in 32x32 pixel greyscale inputimages. The ability to process higher resolution images requires larger and more convolutional layers, so this technique is constrained by the availability of computing resources[21].

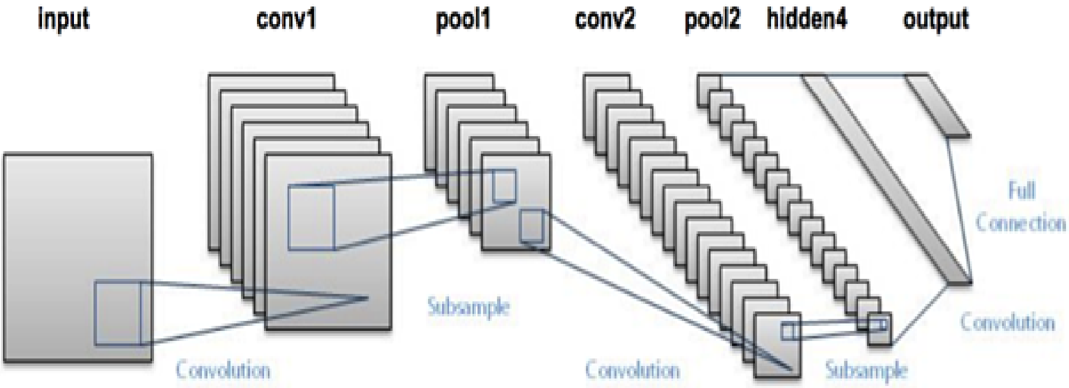


Figure 2.15: le Net architecture

AlexNet :

The first work that popularized the convolutional networks in the vision by ordina- was AlexNet, developed by Alex Krizhevsky, Ilya Sutskever and Geoff Hinton. AlexNet has been challenged by ImageNet ILSVRC [22] in 2012 and significantly outperformed its competitors. The network had an architecture very similar to LeNet, but was deeper, bigger, and had 35 convoluted layers stacked on top of each other (previously, it was common to have only one convolutional layer always immediately followed by a pooling layer)[23].

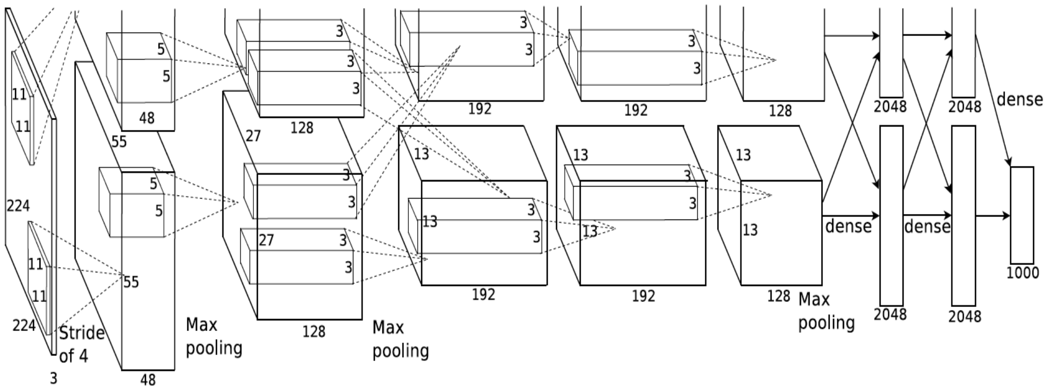


Figure 2.16: AlexNet architecture

VGG Net :

The VGG Network was introduced by the researchers at Visual Graphics Group at Oxford (hence the name VGG). This network is specially characterized by its pyramidal shape, where the bottom layers which are closer to the image are wide, whereas the top layers are deep.[24]

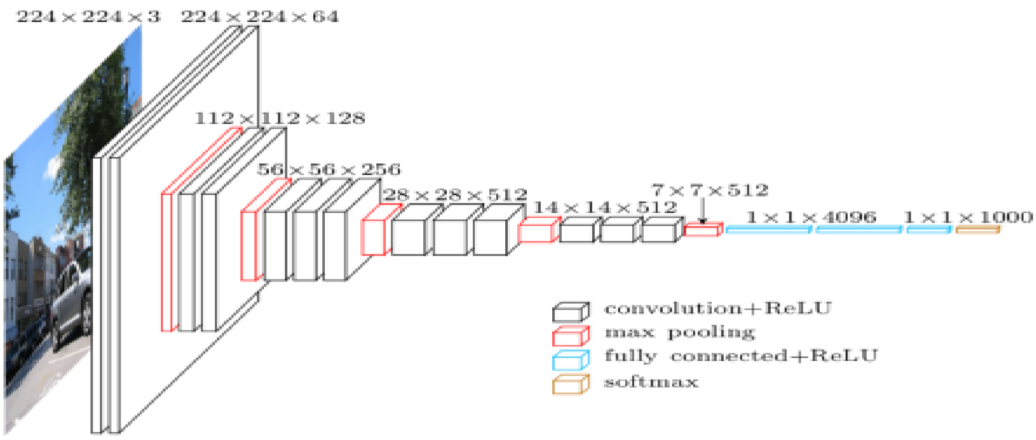


Figure 2.17: VGG-Net architecture

As the image depicts, VGG contains subsequent convolutional layers followed by pooling layers. The pooling layers are responsible for making the layers narrower. In their paper, they proposed multiple such types of networks, with change in deepness of the architecture.

U-net :

The UNET was developed by Olaf Ronneberger et al. for Bio Medical Image Segmentation. The architecture contains two paths. First path is the contraction path (also called as the encoder) which is used to capture the context in the image. The encoder is just a traditional stack of convolutional and max pooling layers. The second path is the symmetric expanding path (also called as the decoder) which is used to enable precise localization using transposed convolutions. Thus it is an end-to-end fully convolutional network (FCN)[25].

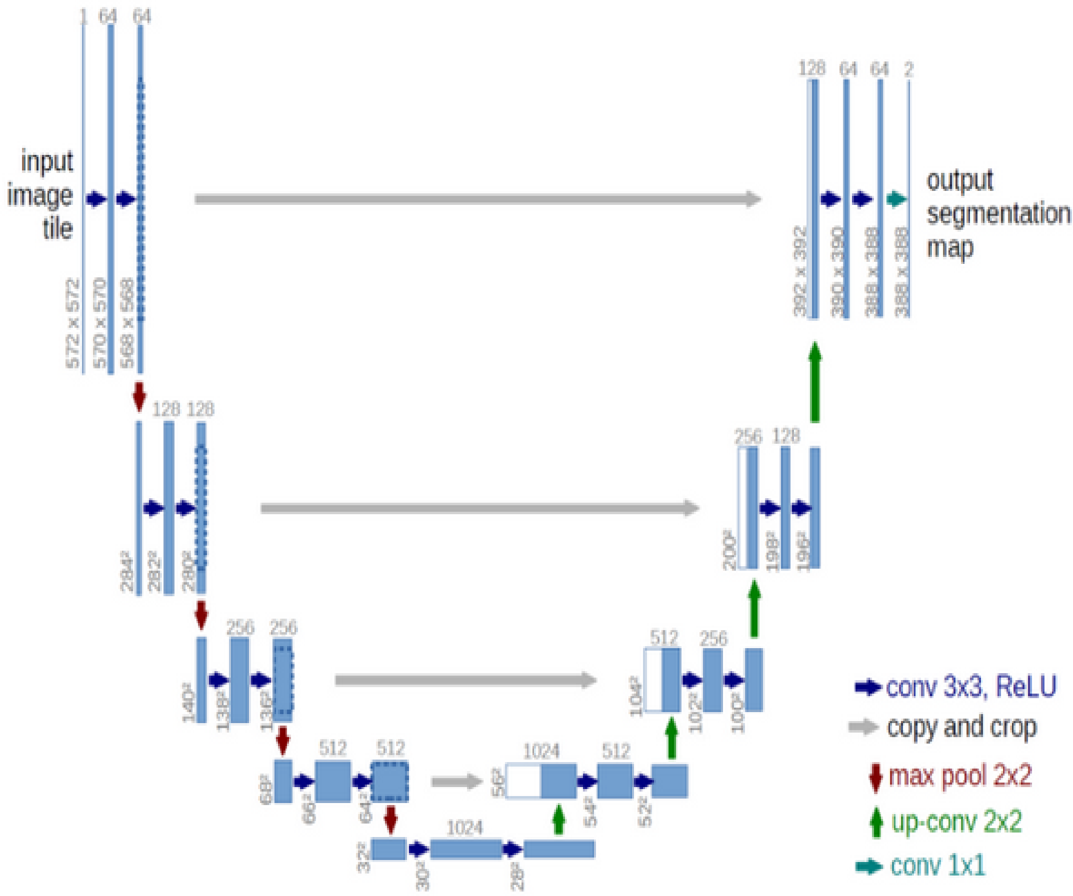


Figure 2.18: U-Net architecture

ResNet :

At last, at the ILSVRC 2015, the so-called Residual Neural Network (ResNet) by Kaiming He et al introduced anovel architecture with "skip connections" and features heavy batch normalization. Such skip connections are also known as gated units or gated recurrent units and have a strong similarity to recent successful elements applied in RNNs. Thanks to this technique they were able to train a Neural Network with 152 layers while still having lower complexity than VGGNet. It achieves a top-5 error rate of 3.57 % which beats human-level performance on this dataset.[23]

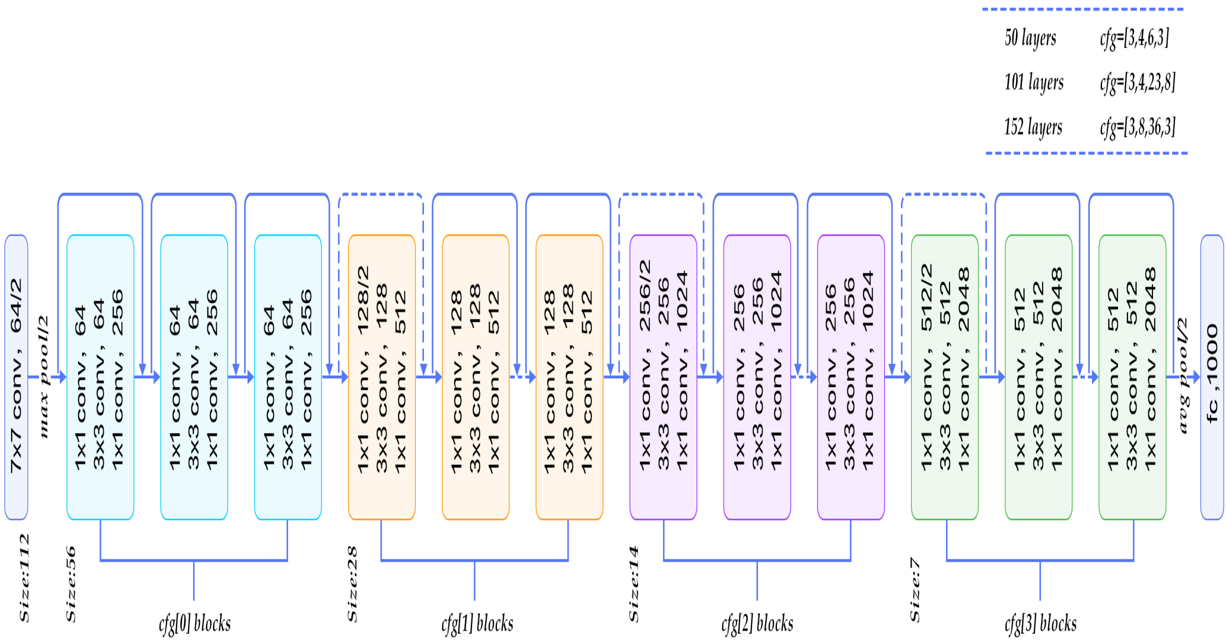


Figure 2.19: ResNet architecture

2.4 Conclusion

In this part, we saw what deep learning is and its relationship with Traditional ML algorithms, We also briefly explained the principle of work with a simplified example, finally we introduced some famous convolution networks and we explained the principle of each. In the next chapter we will use U-net network as architecture for our deep learning model, our choice is motivated by the efficiency of this architecture in segmentation and classification of salt structure in seismic data .

CHAPTER 3

DEEP LEARNING AND BIG DATA

3.1 Introduction

Deep Learning and Big Data Analysis are two focal points of data science. Deep learning models had a great successful results in many applications at artificial intelligence such as computer vision, speech recognition, human emotion analysis, object detection and segmentation, and many other useful applications, especially in health care ...etc

One of the most important areas of using deep learning is analyzing a massive amount of data (Big Data) witch is important for organizations that need to collect a huge amount of data, Deep Learning can be used to extract incredible information that buried in a Big Data.

In this chapter, we will provide a brief survey on some of the most popular deep learning methods, then we will explore some advantages of using Deep Learning in Big Data analysis, including the application of Deep Learning algorithms for Big Data analysis and how specific characteristics of Big Data can lead to some challenges in adopting Deep Learning algorithms for Big Data analytics tasks.

3.2 Big data

The *Big Data* term generally refers to data that exceeds the typical storage , processing, and computing capacity of conventional databases and data analysis techniques, it was defined by Gartner in 2012 [26] as follow "Big Data is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision-making, and process automation"

3.3 Big data sources

Nowadays, big data is everywhere and it can help organisations any industry in many different ways. But what are the various sources of big data ?
In this section we present some of top sources of big data [27] :

Media is the most popular source of big data, as it provides valuable insights on consumer preferences and changing trends. Since it is self-broadcasted and crosses all physical and demographical barriers, it is the fastest way for businesses to get an in-depth overview of their target audience, draw patterns and conclusions, and enhance their decision-making. Media includes social media and interactive platforms, like Google, Facebook, Twitter, YouTube, Instagram, as well as generic media like images, videos, audios, and podcasts that provide quantitative and qualitative insights on every aspect of user interaction.

Cloud

Today, companies have moved ahead of traditional data sources by shifting their data on the cloud. Cloud storage accommodates structured and unstructured data and provides business with real-time information and on-demand insights. The main attribute of cloud computing is its flexibility and scalability [1]. As big data can be stored and

[1]Scalability is the property of a system to adapt to a growing amount of work by adding resources to the system.

sourced on public or private clouds, via networks and servers, cloud makes for an efficient and economical data source.

Internet of things (IoT)

Machine-generated content or data created from IoT constitute a valuable source of big data. This data is usually generated from the sensors that are connected to electronic devices. The sourcing capacity depends on the ability of the sensors to provide real-time accurate information. IoT is now gaining momentum and includes big data generated, not only from computers and smartphones, but also possibly from every device that can emit data. With IoT, data can now be sourced from medical devices, vehicular processes, video games, meters, cameras, household appliances, and the like.

3.4 Advantages of Using Deep Learning in Big Data

algorithms of machine learning have solved many problems,but they have failed to solve the biggest problems facing artificial intelligence, such as speech recognition and object recognition, however This failure stimulate researchers to develop them to show what is known as deep learning
With the advent of deep learning, the field of artificial intelligence has developed considerably, But with getting larger amounts of data are available thanks to Big Data and connected objects and the computing machines have become more powerful that we have understood the real power of Deep Learning.This power lies in the ability of the extraction of features witch done automatically by the algorithm, Where a high amounts of data improves the results of this process.
Thanks to the entry of the concept of Big data on deep learning this last break into many scientific fields such as learned to drive a car [28] , diagnose cancer [29] and autism[30] over

the last years

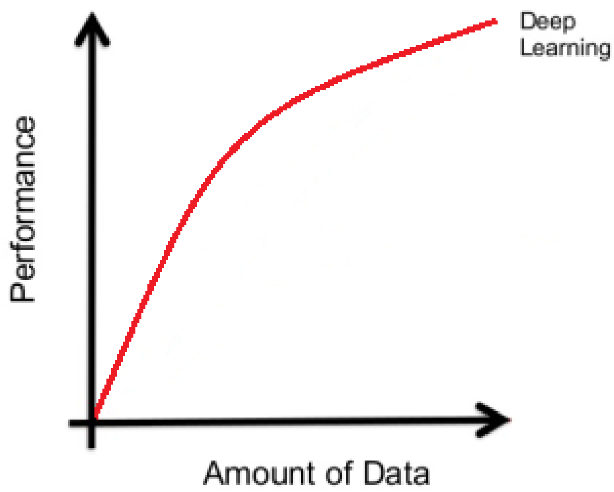


Figure 3.1: Big Data and Deep Learning

3.5 Applications of Deep Learning in Big Data Analytics

Semantic Indexing

Information retrieval is a key task that is associated with Big Data Analytics. Efficient storage and retrieval of information is a growing problem in Big Data analysis, as data in large-scale quantities such as text, image, video, and audio is being collected and made available across various domains. Thus, strategies and solutions that were previously used for information storage and retrieval are challenged by massive volumes of data. Semantic indexing proves to be an efficient technique as it facilitates knowledge discovery and comprehension, thereby making search engines work more quickly and efficiently. [31]

Conducting Discriminative Tasks

While performing discriminative tasks in Big Data Analytics, Learning algorithms allow users to extract complicated nonlinear features from the raw data. It also facilitates the use of linear models to perform discriminative tasks using extracted features as input. This approach has two advantages: Firstly, by extracting features with Deep Learning adds nonlinearity to the data analysis, thereby associating discriminative tasks closely to AI, and secondly applying linear analytical models on extracted features is more efficient computationally. These two benefits are important for Big Data because it allows practitioners in accomplishing complicated tasks related to Artificial Intelligence like object recognition in images, image comprehension, etc. [31]

Semantic Image and Video Tagging

Deep Learning techniques help in semantic tagging. Deep Learning mechanisms can facilitate segmentation and annotation of complex image scenes. Deep Learning can also be used for action scene recognition as well as video data tagging. It uses an independent variant analysis to learn invariant spatiotemporal features from video data. This approach helps in extracting useful features for performing discriminative tasks on image and video data. Deep Learning has been successful in achieving remarkable results in extracting useful features. However, there is still considerable work that remains to be done for further exploration that includes determination of appropriate objectives in learning good data representations and performing other complex tasks in Big Data Analytics.[31]

3.6 Challenges in Adopting Deep Learning Algorithms for Big Data Analytics

The prior section focused on the applicability and benefits of Deep Learning algorithms for Big Data Analytics, On the other hand, Big Data associated with some characteristic that make a challenge in adapting Deep Learning, This section presents some areas of Big Data where Deep Learning needs further exploation, specifically, learning with streaming data, dealing with high-dimensional data.

Incremental learning for non-stationary data

One of the challenging aspects in Big Data Analytic is dealing with streaming and fast moving input data, in some application of artificial intelligent it is important to adapt Deep Learning to handle streaming data, as there is a need for algorithms that can deal with large amounts of continuous input data, in this part we select three paper witch include works associated with Deep Learning and streaming data,including incremental feature learning and extraction[32], denoising autoencoders[33], and deep belief networks[34].

Zhou et al [32] shows that the Deep Learning algorithm can be a solution for incremental feature learning on very large datasets with employing denoising autoencoders [33] witch are a variant of autoencoders which extract features from corrupted input, where the extracted features are robust to noisy data and good for classification purposes.

Calandra et al. [34] introduce adaptive deep belief networks which demonstrates how Deep Learning can be generalized to learn from online non-stationary and streaming data. Their study exploits the generative property of deep belief networks to mimic the samples from the original data, where these samples and the new observed samples are used to learn the new deep belief network which has adapted to the newly observed data.

High-dimensional data

Some Deep Learning algorithms can become prohibitively computationally-expensive when dealing with high-dimensional data, such as images, likely due to the often slow learning process associated with a deep layered hierarchy of learning data abstractions and representations from a lower-level layer to a higher-level layer. That is to say, these Deep Learning algorithms can be stymied when working with Big Data that exhibits large Volume, one of the four Vs associated with Big Data Analytics. A high-dimensional data source contributes heavily to the volume of the raw data, in addition to complicating learning from the data.[35]

As solution for that problem, Chen et al. [36] introduce marginalized stacked denoising autoencoders (mSDAs) which scale effectively for high-dimensional data and is computationally faster than reg- ular stacked denoising autoencoders (SDAs).

Convolutional neural networks are another method which scales up effectively on high- dimensional data. Researchers have taken advantages of convolutional neural networks on ImageNet dataset with 256x256 RGB images to achieve state of the art results [37, 38]. In convolutional neural networks, the neurons in the hidden layers units do not need to be connected to all of the nodes in the previous layer, but just to the neurons that are in the same spatial area. Moreover, the resolution of the image data is also reduced when moving toward higher layers in the network.

3.7 Conclusion

At the end of this chapter we conclude that mining and extracting meaningful patterns from large data sets for decision-making, and prediction are critical aspects of Big Data analytic. But, Big Data Analytic is associated with some challenges during data mining and extraction, and we mention that Deep Learning algorithms can address these challenges.

CHAPTER 4

PROPOSED METHOD

4.1 Introduction

After having studied in the previous chapters the state of the art of different tools of Deep Learning and Big Data and studying some of recent work on seismic analysis, we will present in this part the proposed architecture for salt classification problem.

4.2 System Architecture

The system proposed consists of two layers illustrate on the figure 4.1

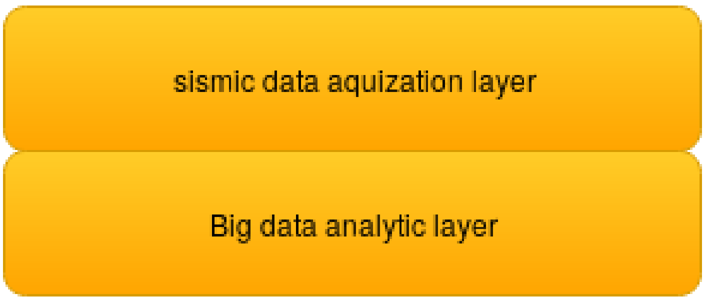


Figure 4.1: layer of the system

The first layer focus on the geological treatment for getting a useful data, these processes can be expressed by figure 4.3 :

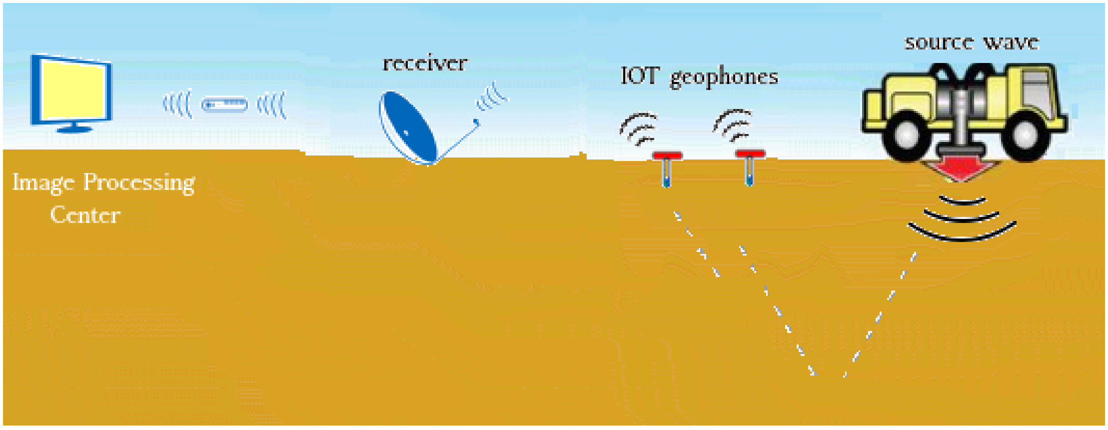


Figure 4.2: Seismic data acquisition layer

The second layer concern the storage of the data and deep learning model, we had used the Hadoop. It is a free and open source framework written in Java to facilitate the creation of distributed and scalable applications that allow applications to work with thousands of nodes and petabytes of data.

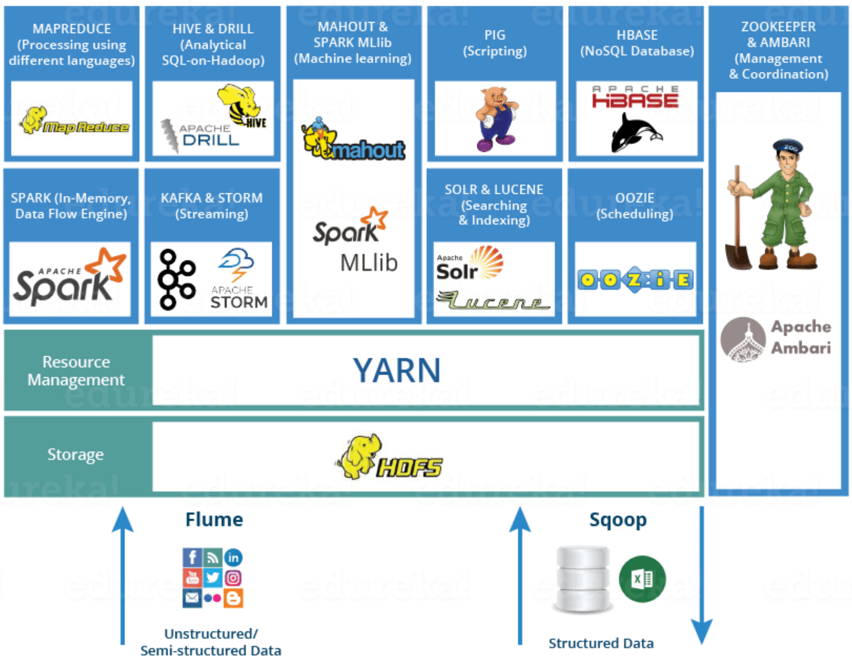


Figure 4.3: Hadoop components

4.3 Deep Learning Architecture

The proposed architecture of the neural network is based on the U-Net architecture. We worked to improve the original architecture witch explained previously in chapter tow. So by adding a batch normalization layer after each convolution layer and increasing the depth with a new block at the bottom of U-Net.

Batch normalization is a method that we can use to normalize the inputs of each layer, in order to fight the internal covariate shift problem. This problem appear when we do some preprocessing to the input data. For example, we could normalize all data so that it resembles to a normal distribution (that means, zero mean and a unitary variance), some reasons for that preprocessing are : preventing the early saturation of non-linear activation functions like the sigmoid function, assuring that all input data is in the same range of values, etc.

During training time, a batch normalization layer does the following:

1. Calculate the mean and variance of the layers input.

$$\mu_B = \frac{1}{m}\Sigma_{i=1}^{m} x_i \qquad \textit{Batch mean}$$

$$\sigma_B^2 = \frac{1}{m}\Sigma_{i=1}^{m} (x_i - \mu_B)^2 \qquad \textit{Batch variance}$$

2. Normalize the layer inputs using the previously calculated batch statistics.

$$\overline{x_i} = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}$$

3. Scale and shift in order to obtain the output of the layer.

$$y_i = \gamma \overline{x_i} + \beta$$

In order to add a batch normalization layer in our model, Tensorflow API already has all this mathematical formulas implemented . We just need to import the batch normalization package with the following code:

```
from keras.layers import BatchNormalization
```

and type three lines of code bellow after each convolution layer

```
bn = BatchNormalization()
n  = BatchNormalization()(convLayer) if bn else convLayer
n  = Dropout(0)(n) if 0 else n
```

We can illustrate the architecture proposed with the figure 4.4

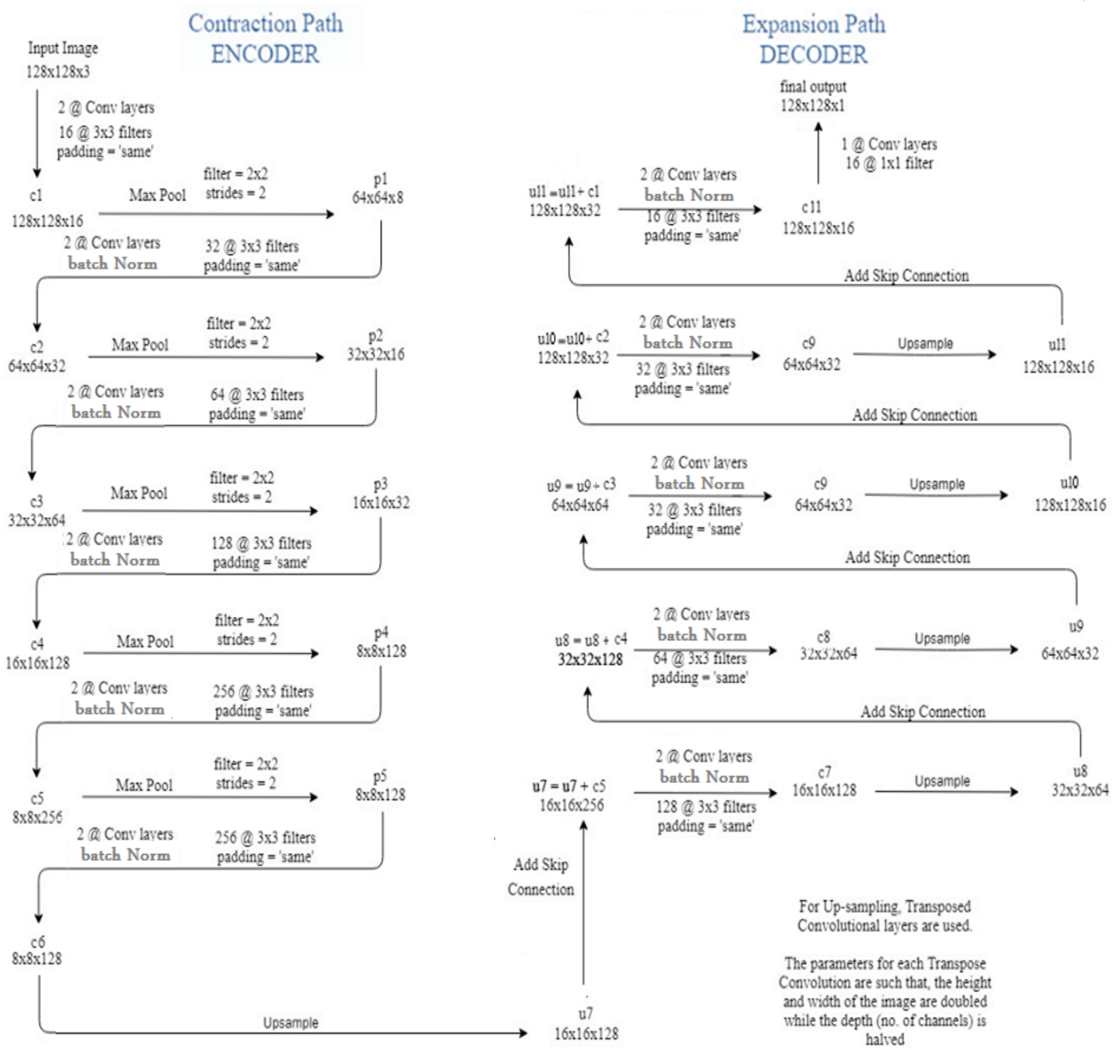


Figure 4.4: Deep Learning architecture based on U-net architecture

4.4 Used Data

The data-set used in this project is provided by geoscience data company (TGS), TGS is committed to providing the global energy industry with the right subsurface data, in the right place, at the right time, so that the customers can make informed drilling decisions. the data is a set of images chosen at various locations chosen at random in the subsurface. The images are 101 x 101 pixels and each pixel is classified as either salt or sediment. In addition to the seismic images, the depth of the imaged location is provided for each image. The data-set is divided into two parts,The training set has 4000 x 2 image includes both the image and salt-segmentation mask.The test set include 18000 simple for prediction. the figure billow give visualization to the used data to get better understanding:

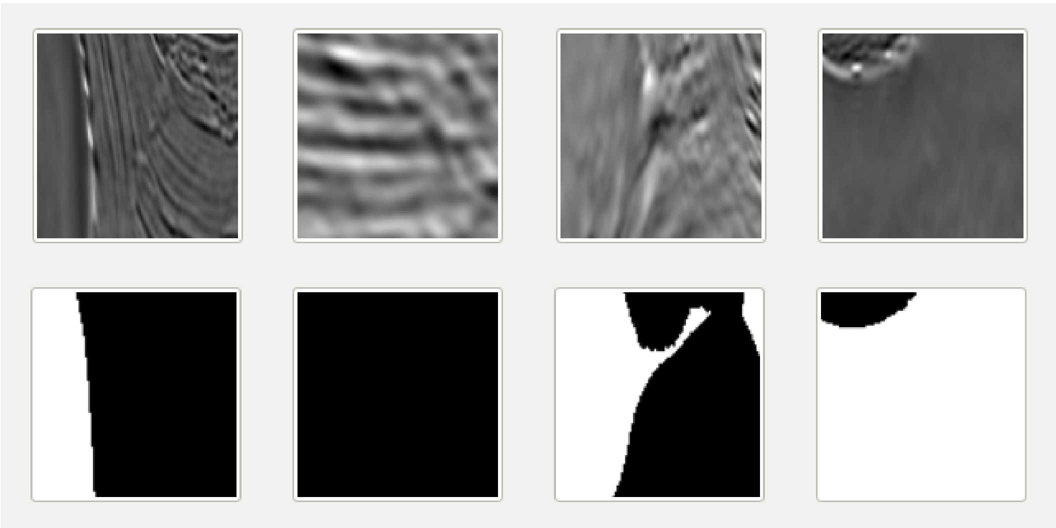


Figure 4.5: Training samples from TGS data-set

Notice that if the mask is entirely black, this means there are no salt deposits in the given seismic image.
Clearly from the above few images it can be inferred that its not easy for human experts to make accurate mask predictions for the seismic images.

4.5 Result

The result of this work is evaluated on the mean average precision at different intersection over union (IoU) thresholds. The IoU of a proposed prediction and a set of true prediction calculated as:

$$IoU = \frac{A \cap B}{A \cup B}$$

the implementation of this function shown in the figure 4.6

```
# Define IoU metric
def mean_iou(y_true, y_pred):
    prec = []
    for t in np.arange(0.5, 1.0, 0.05):
        y_pred_ = tf.to_int32(y_pred > t)
        score, up_opt = tf.metrics.mean_iou(y_true, y_pred_, 2)
        K.get_session().run(tf.local_variables_initializer())
        with tf.control_dependencies([up_opt]):
            score = tf.identity(score)
        prec.append(score)
    return K.mean(K.stack(prec), axis=0)
```

Figure 4.6: IoU implementation

Lastly, the score returned by the metric evaluation function After 10 epochs in machine i5 with ram 6 GB and in less then 1 hour of training is 68% .

4.6 Conclusion

In this part, we described the approach taken to obtain the classifier of 2D seismic image, We started with the original U-net architecture and we worked to improve it whether on the level of prediction or training time, also we explore the used data and storage architecture in Hadoop, lastly we present the obtained result on the metric evaluation function.
We conclude that the additions (batch normalization layers, increasing the depth, big data architecture) can raise the prediction ratio from 62% in 30 epoch to 68% in 10 epoch.

CONCLUSION AND PERSPECTIVES

In this study, we explored the field of image classification that like all the others fields of artificial intelligence has undergone a major evolution since the appearance of the Deep Learning.

We spent a lot of time reading publications and articles in relation to deep learning and big data to see what is best about classification and to be able to design our own model. This work allowed us to put into practice our knowledge about neural networks.

IOU was increased to 68% with the following steps :

- Added a batch normalization layer after each convolutional layer.
- Increased the depth to 5 with a new layer at the bottom of U-Net

As perspectives we can mention:

- Implementing the techniques that make the success of models that participate in the Imagenet challenge (Vgg-Net, Alex-net, etc...).
- Segmentation and generation of images.

REFERENCES

[1] N. Mondol and K. Bjørlykke, *Seismic Exploration*, 09 2010, pp. 375–402.

[2] M. P. A. Jackson and M. R. Hudec, *Seismic Interpretation of Salt Structures.* Cambridge University Press, 2017, p. 364–398.

[3] A. Francis, "A Simple Guide to Seismic Inversion," https://www.geoexpro.com/articles/2014/06/a-simple-guide-to-seismic-inversion, 2014.

[4] C. P. North, "Gluyas, j & swarbrick, r., 2004 petroleum geoscience, blackwell science ltd," *Teaching in Earth Sciences*, vol. 28, pp. 41–42, 2003.

[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[6] G. Hinton, L. Deng, D. Yu, G. Dahl, A.-r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, B. Kingsbury *et al.*, "Deep neural networks for acoustic modeling in speech recognition," *IEEE Signal processing magazine*, vol. 29, 2012.

[7] O. Lavialle, S. Pop, C. Germain, M. Donias, S. Guillon, N. Keskes, and Y. Berthoumieu, "Seismic fault preserving diffusion," *Journal of Applied Geophysics*, vol. 61, no. 2, pp. 132–141, 2007.

[8] I. Pitas and C. Kotropoulos, "Texture analysis and segmentation of seismic images," in *International Conference on Acoustics, Speech, and Signal Processing,.* IEEE, 1989, pp. 1437–1440.

[9] M. Karchevskiy, I. Ashrapov, and L. Kozinkin, "Automatic salt deposits segmentation: A deep learning approach," *arXiv preprint arXiv:1812.01429*, 2018.

[10] A. Sahin, "Anseismic – a simple assistant computer program for implementing earthquake analysis of structures with ansys and sap2000," 09 2010.

[11] M. D. Youcef, "Deep learning pour la classification des images," Ph.D. dissertation, 2016-2017.

[12] What is machine learning? a definition. [Online]. Available: https://www.expertsystem.com/machine-learning-definition/

[13] 9 algorithmes de machine learning que chaque data scientist doit connaitre. [Online]. Available: https://mrmint.fr/9-algorithmes-de-machine-learning-que-chaque-data-scientist-doit-connaitre

[14] K. Hechenbichler and K. Schliep, "Weighted k-nearest-neighbor techniques and ordinal classification," 2004.

[15] I. Rish *et al.*, "An empirical study of the naive bayes classifier," in *IJCAI 2001 workshop on empirical methods in artificial intelligence*, vol. 3, no. 22, 2001, pp. 41–46.

[16] J. A. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural processing letters*, vol. 9, no. 3, pp. 293–300, 1999.

[17] R. K. Mehra and J. Peschon, "An innovations approach to fault detection and diagnosis in dynamic systems," *Automatica*, vol. 7, no. 5, pp. 637–640, 1971.

[18] Understanding k-means clustering in machine learning. [Online]. Available: https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-6a6e67336aa1

[19] D. P. Mandic and J. Chambers, *Recurrent neural networks for prediction: learning algorithms, architectures and stability.* John Wiley & Sons, Inc., 2001.

[20] I. Aizenberg, N. N. Aizenberg, and J. P. Vandewalle, *Multi-Valued and Universal Binary Neurons: Theory, Learning and Applications.* Springer Science & Business Media, 2013.

[21] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner *et al.*, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[22] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[23] Cnn architectures: Lenet, alexnet, vgg, googlenet, resnet and more .... [Online]. Available: https://medium.com/@sidereal/cnns-architectures-lenet-alexnet-vgg-googlenet-resnet-and-more-666091488df5

[24] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[25] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention.* Springer, 2015, pp. 234–241.

[26] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," *International journal of information management*, vol. 35, no. 2, pp. 137–144, 2015.

[27] Top 5 sources of big data. [Online]. Available: https://www.allerin.com/blog/top-5-sources-of-big-data

[28] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang *et al.*, "End to end learning for self-driving cars," *arXiv preprint arXiv:1604.07316*, 2016.

[29] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, p. 115, 2017.

[30] H. C. Hazlett, H. Gu, B. C. Munsell, S. H. Kim, M. Styner, J. J. Wolff, J. T. Elison, M. R. Swanson, H. Zhu, K. N. Botteron *et al.*, "Early brain development in infants at high risk for autism spectrum disorder," *Nature*, vol. 542, no. 7641, p. 348, 2017.

[31] 3 applications of deep learning in big data analytics. [Online]. Available: https://www.allerin.com/blog/3-applications-of-deep-learning-in-big-data-analytics

[32] G. Zhou, K. Sohn, and H. Lee, "Online incremental feature learning with denoising autoencoders," in *Artificial intelligence and statistics*, 2012, pp. 1453–1461.

[33] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proceedings of the 25th international conference on Machine learning.* ACM, 2008, pp. 1096–1103.

[34] R. Calandra, T. Raiko, M. P. Deisenroth, and F. M. Pouzols, "Learning deep belief networks from non-stationary streams," in *International Conference on Artificial Neural Networks.* Springer, 2012, pp. 379–386.

[35] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *Journal of Big Data*, vol. 2, no. 1, p. 1, 2015.

[36] M. Chen, Z. Xu, K. Weinberger, and F. Sha, "Marginalized denoising autoencoders for domain adaptation," *arXiv preprint arXiv:1206.4683*, 2012.

[37] G. E. Dahl, D. Yu, L. Deng, and A. Acero, "Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition," *IEEE Transactions on audio, speech, and language processing*, vol. 20, no. 1, pp. 30–42, 2011.

[38] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.